

# Bilingual Sentiment Embeddings: Joint Projection of Sentiment Across Languages

Jeremy Barnes<sup>1,2</sup>, Roman Klinger<sup>1</sup>, Sabine Schulte im Walde<sup>1</sup>  
University of Stuttgart<sup>1</sup>, Universitat Pompeu Fabra<sup>2</sup>



## Motivation

**Goal:** How can you create bilingual word representations for cross-lingual sentiment analysis when you have little parallel data?

**Problem:** There is no labeled sentiment data or large amounts of parallel data available in most languages in the world.

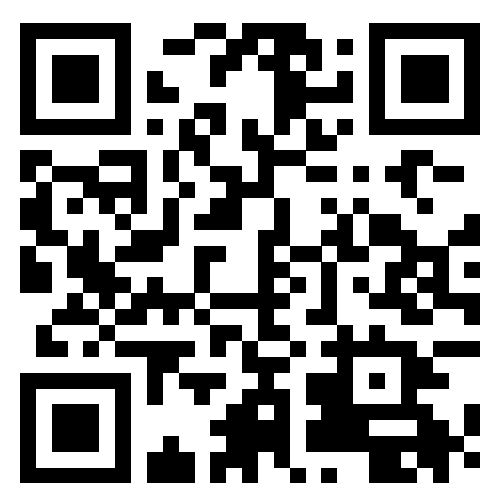
**Approach:** We jointly train the projection and sentiment classifier.

## Experimental Setup

- *Sentence-level* Cross-lingual Sentiment Analysis on hotel reviews: the *OpenNER* and *MultiBooked Corpora*. This avoids additional domain problems.
- *Three language pairs:*
  - English → Spanish (ES)
  - English → Catalan (CA)
  - English → Basque (EU)
- *Binary and 4-class.*
- Train on the English annotated sentiment corpus only. Test on the target language test set.
- Compare to two bilingual embedding methods with similar data requirements (ARTETXE, BARISTA) and Machine Translation (MT).

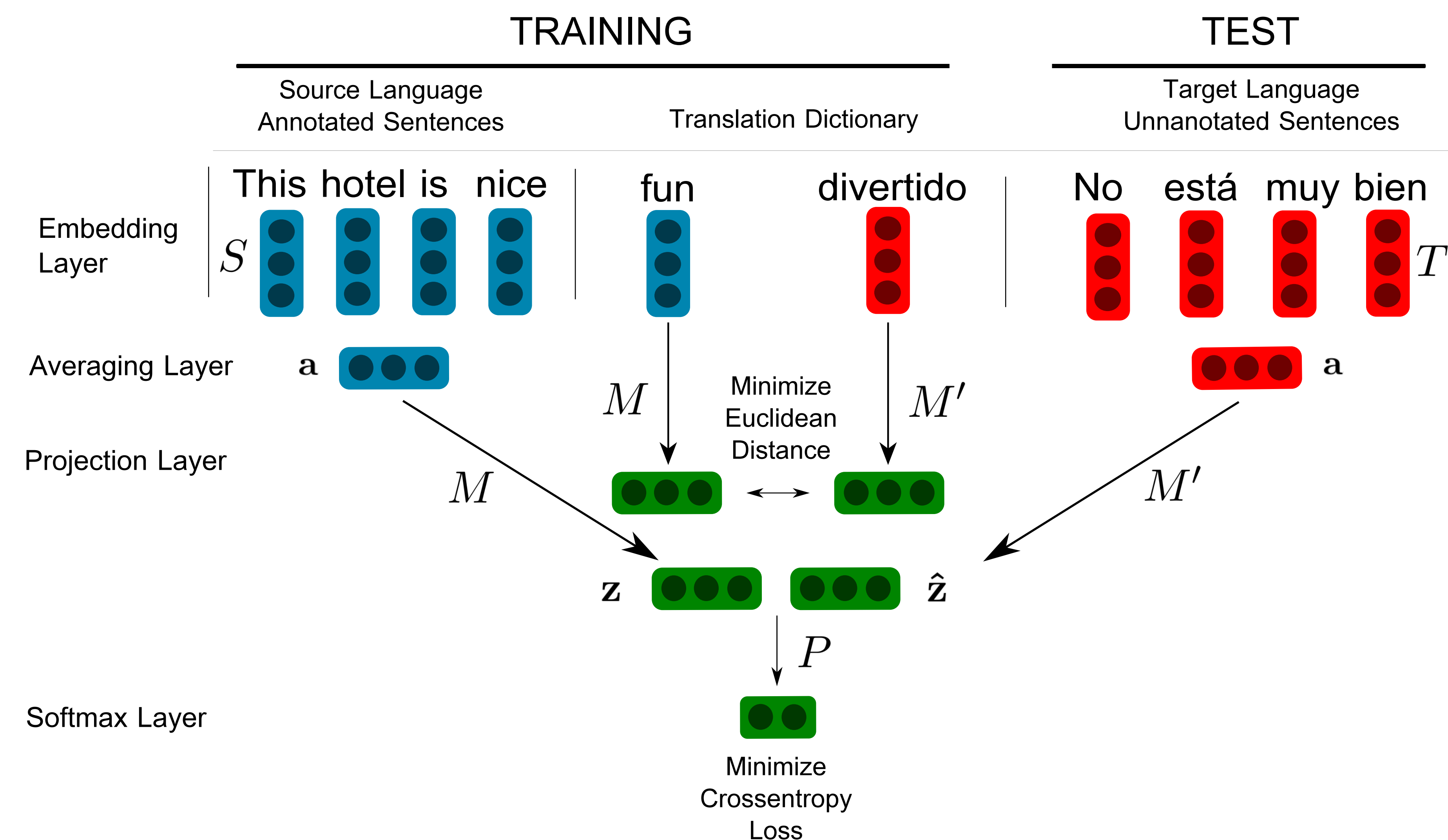
## Code Available

<https://github.com/jbarnesspain/blse>



Contact: [jeremy.barnes@upf.edu](mailto:jeremy.barnes@upf.edu)

## Bilingual Sentiment Embedding Model



## Results for Cross-lingual Sentiment Classification

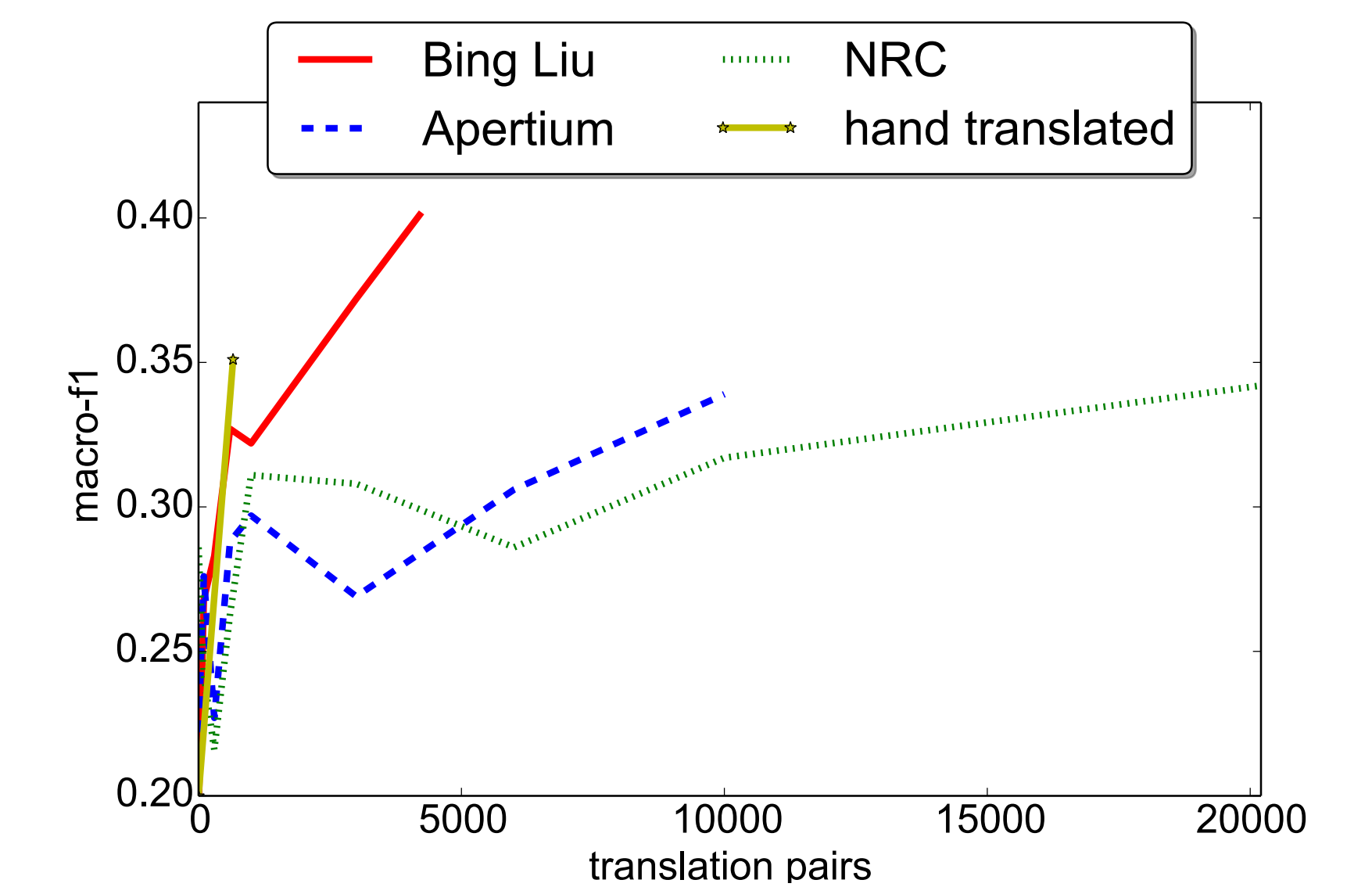
- BLSE *outperforms other projection methods* on macro  $F_1$  ( $p < 0.05$  highlighted) and is close to MT.
- ENSEMBLE of BLSE and MT performed better than other combinations (not shown here).

		ARTETXE	BARISTA	BLSE	MT	ENSEMBLE
Binary	ES	67.1	61.2	74.1	79.0	80.3
	CA	60.7	60.1	72.9	77.2	85.0
	EU	45.6	54.8	69.3	69.4	73.5
4-class	ES	34.9	39.5	41.2	48.8	50.3
	CA	23.0	36.2	35.9	53.7	53.9
	EU	21.3	33.8	30.0	43.6	50.5

## Projection Lexicon

Translating a sentiment lexicon is better than using a general bilingual dictionary.

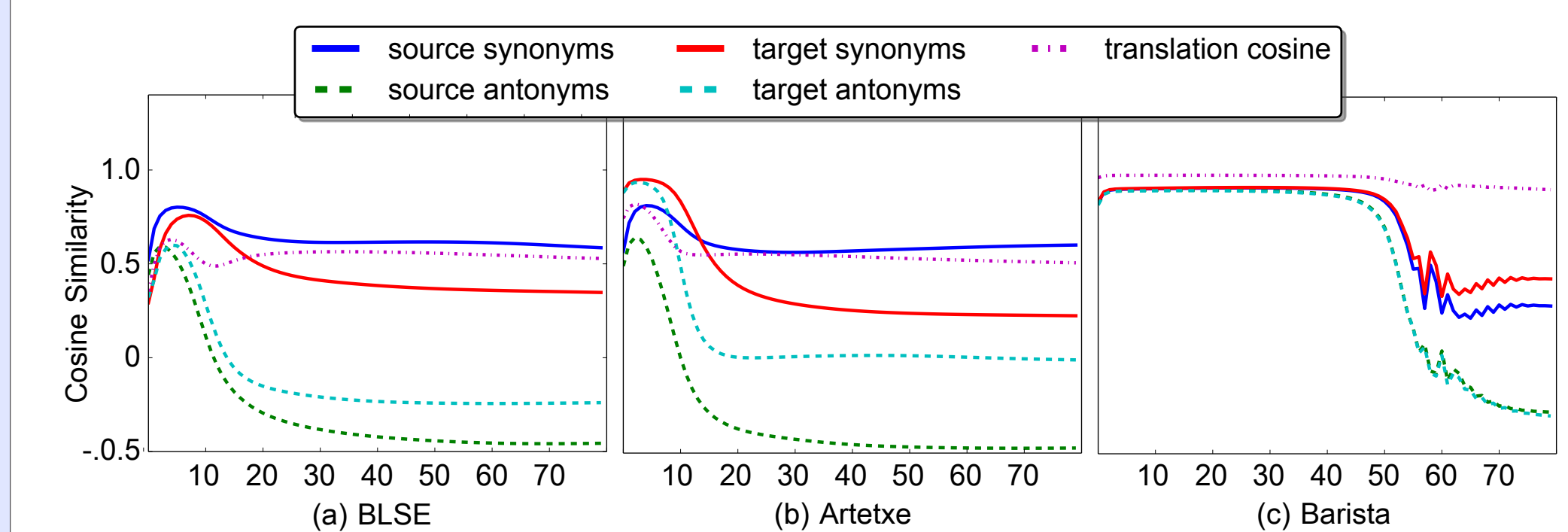
Even a small hand-translated dictionary (~600 pairs) can outperform large general dictionaries.



## Model Analysis

BLSE learns that "sentiment synonyms" should be close in vector space and "sentiment antonyms" should be farther away.

BLSE performs better on binary task. It (1) assigns too much sentiment to functional words and (2) does not preserve the original semantics of the embeddings.



## Future Work

- Move our approach to aspect-level.
- Include multi-word phrases into projection to improve negation and modifiers.